

DATABASE SYSTEM INCLUDING CENTER SERVER
AND LOCAL SERVERS

BACKGROUND OF THE INVENTION

The present invention relates to a database system and a method for accessing a center server and a database and more particularly, to a database system
5 for replicating databases in a single or a plurality of storage subsystems which are connected to a network and located remotely and permitting replicated databases to be accessed in a consolidated fashion and a method for accessing a center server and databases in that system.

10 As a conventional technique concerning a method for replicating a remotely located database, a technique of database replication has been known in which data is replicated between servers mutually connected through a LAN or WAN. As another con-
15 ventional technique, a technique of database hub has been known according to which inquiries are made to distributed database management systems during execution of the inquiries and results returned from the individual database management systems are consolidated
20 so as to be exhibited as one result. Further, known as still another conventional technique directed to preparation of a replication of a database at a remote location is a technique of disaster recovery using a volume replication function owned by a storage device.

25 Known as a prior art concerning the database

hub is a technique described in "Data Joiner",
International Business Machines Corporation, Internet
document searched on November 11, 2002,
<URL:http://www-3.ibm.com/software/data/datajoiner/>
5 and known as prior arts concerning the disaster
recovery are techniques described in "Building Storage
Networks" by Marc Farley, second edition, Osbone/
McGraw-Hill Corp., 2001, pp. 118-124 and described in
"HiRDB version 6", Hitachi Com. Software Department,
10 Internet document searched on November 11, 2002,
<URL:http://www.hitachi.co.jp/Prod/comp/soft/open-
e/hirdb/v6/outline/confv6.htm>.

When a method based on the aforementioned
database replication is used as means for getting a
15 consolidated access to a plurality of databases at
remote locations, the LAN or WAN is used for transfer
of data, raising a problem that much time is consumed
for replication of data. In a method using the
database hub, the remotely located database management
20 systems are accessed during execution of inquiries,
with the result that the response time is degraded and
besides, when a large number of results are brought
about, a large amount of data must be transferred
through the medium of the LAN or WAN, giving rise to a
25 problem that the search performance is degraded. In
addition, the method of disaster recovery using the
volume replication function the storage unit has is
back-up for a database on the replication side to

recover and disadvantageously, it cannot afford to consolidate a plurality of databases.

SUMMARY OF THE INVENTION

It is an object of the present invention to
5 provide a database system for permitting a plurality of databases at remote locations to be accessed instantaneously in a consolidated fashion and a method for accessing a center server and databases in the system.

10 According to the invention, to accomplish the above object, in a database system comprising a center server, a single or a plurality of local servers, a first network for mutually connecting the center server and the local servers, local storage subsystems for
15 storing local databases managed by the local servers, a center storage subsystem for storing replications of the local databases and a second network for mutually connecting the center server, the center storage subsystem, the local servers and the local storage
20 subsystems, the center server includes a replication requesting unit for requesting the local servers to replicate local databases and a data consolidating unit for performing a process of consolidating the replicated local databases, and each of the local
25 servers includes a local database freeze requesting unit responsive to a database replication request from the center server to request a database management

system to freeze the local database and a database replicating unit for causing the local storage subsystem to replicate, in the center storage subsystem, the local database the local storage subsystem stores.

5 Also to accomplish the above object, in a method for accessing a database system comprising a center server, a single or a plurality of local servers, a first network for mutually connecting the center server and the local servers, local storage subsystems
10 for storing local databases managed by the local servers and a second network for mutually connecting the center server, the center storage subsystem, the local servers and the local storage subsystems, the center server requests the local servers to replicate
15 the local databases and performs a process of consolidating the replicated local databases, and each of the local servers responds to a database replication request from the center server to request a database management system to freeze the local database and
20 causes the local storage subsystem to replicate, in the center storage subsystem, the local database the local storage subsystem stores.

Other objects, features and advantages of the invention will become apparent from the following
25 description of the embodiments of the invention taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block diagram showing the construction of a database system according to an embodiment of the invention.

5 Fig. 2 is a flowchart for explaining a process in which synchronization between each shadow image and each local DB is set up and data of a plurality of local DB's is accessed in a consolidated fashion through a center DBMS.

10 Fig. 3 is a table for explaining an example of structure of a replication source managing table provided in a center server.

 Fig. 4 is a flowchart for explaining a process executed by a data consolidating unit in the
15 center server (a process in step 407 explained in connection with Fig. 2).

 Fig. 5 is a table showing an example of the replication source managing table after completion of synchronization of only a DB at area A.

20 Fig. 6 is a table showing an example of the replication source managing table after completion of synchronization of two DB's at areas A and B.

DESCRIPTION OF THE EMBODIMENTS

Embodiments of database system and method for
25 accessing databases according to the invention will be described hereunder in greater detail with reference to the drawings.

Fig. 1 is a block diagram showing the construction of a database system according to an embodiment of the invention. In Fig. 1, reference numeral 101 designates a center server, 102a and 102b
5 local servers, 103 a center storage subsystem, 104a and 104b local storage subsystems, 105 a LAN, 106 a SAN (storage area network), 110 a replication requesting unit, 112 a data consolidation completion notifying unit, 2 a replication source managing table, 3 a data
10 consolidating unit, 14, 129a and 129b database management systems (BDMS's), 115 a volume synchronizing unit, 116 a volume splitting unit, 117 a volume replicating unit, 121a and 121b DB (database) freeze requesting units, 122a and 122b remote volume split
15 request units, 123a and 123b remote volume split completion notifying units, 124a and 124b DB freeze release requesting units, 125a and 125b remote volume resynchronization requesting units, 126a and 126b remote volume splitting units, 127a and 127b remote
20 volume resynchronizing units, 128a and 128b remote volume replicating units, 131a and 131b local shadow images, 133a and 133b replication local DB's, and 134a and 134b local DB's.

The database system according to one
25 embodiment of the invention comprises, as shown in Fig. 1, the center server 101, local servers 102a and 102b provided in a plurality areas A and B, center storage system 103, local storage subsystems 104a and 104b

provided in the plurality of areas A and B, LAN 105 serving as the first network for connecting the center server 101 with the local servers 102a and 102b, and SAN 106 serving as the second network for connecting
5 the center server 101 with the local servers 102a and 102b, center storage subsystem 103 and local storage subsystems 104a and 104b.

In the above construction, the center server 101 includes replication requesting unit 110, data
10 consolidating unit 3, data consolidation completion notifying unit 112, replication source managing table 2 and database management system (DBMS) 14. Then, the center storage subsystem 103 includes volume synchronizing unit 115, volume splitting unit 116,
15 volume replicating unit 117, replication local databases 133a and 133b representing replications of the local databases 134a and 134b and shadow images 131a and 131b of the replication local databases 133a and 133b.

20 Each of the local servers 102a and 102b includes local database freeze requesting unit 121a or 121b, remote volume split requesting unit 122a or 122b, remote volume split completion notifying unit 123a or 123b, local DB freeze release requesting unit 124a or
25 124b, remote volume resynchronization requesting unit 125a or 125b and local DBMS 129a or 129b. Then, each of the local storage subsystems 104a and 104b includes remote volume splitting unit 126a or 126b, remote

volume resynchronizing unit 127a or 127b, remote volume replicating unit 128a or 128b and local DB 134a or 134b.

The local servers 102a and 102b are provided in the areas A and B, respectively, and the local
5 DBMS's 129a and 129b incorporated in the respective local servers manage the local DB's 134a and 134b, respectively, which are provided in the corresponding local storage subsystems 104a and 104b, respectively.

In the DB system according to the embodiment
10 of the invention constructed as above, information for updates applied to the local DB's of local storage subsystems 104a and 104b (hereinafter referred to as update information) is transmitted asynchronously to the center storage subsystem 103 by means of the remote
15 volume replicating units 128a and 128b, respectively. The volume replicating unit 117 of the center storage subsystem 103 receiving the update information asynchronously reflects the received update information upon the replication local DB's 133a and 133b. In the
20 case of an example shown in Fig. 1, an update applied to the local DB 134a in the area A is reflected upon the replication local DB 133a and an update applied to the local DB 134b in the area B is reflected upon the replication local DB 133b. Synchronization between
25 local DB 134a and replication local DB 133a and that between local DB 134b and replication local DB 133b are set up by causing the remote volume splitting units 126a and 126b to stop the reflection of the update

information.

The center DBMS 14 incorporated in the center server 101 manages the local shadow images 131a and 131b incorporated in the center storage subsystem 103.

5 These shadow images 131a and 131b are those of the replication local DB's 133a and 133b, respectively. In the case of the example shown in Fig. 1, the center DBMS 14 manages the two shadow images of local shadow images 131a and 131b. The volume synchronizing unit
10 115 of the center storage subsystem 103 reflects the update information, which has been reflected upon the replication local DB's 133a and 133b, upon the local shadow images 131a and 131b.

In the embodiment of the invention described
15 previously, the center server and each local server respectively include a CPU, a memory, a disk device and a driver for external storage medium, which are not shown and are interconnected through a communication path such as a bus. Also, the center storage subsystem
20 and each local storage subsystem respectively include, in addition to the above components, a DB or DB's held in the disk device. Then, the individual functional components constituting the aforementioned servers and storage subsystems are implemented when a program
25 stored in the disk device or external storage medium is written to the memory and executed by the CPU. It is to be noted that only two sets of local servers and local storage subsystems are shown in Fig. 1 but they

may be provided in more increased number.

Fig. 2 is a flowchart for explaining a process in which synchronization between the shadow image 131a managed by the center DBMS 14 and the local DB 134a managed by the local DBMS 129a and that between the shadow image 131b managed by the center DBMS 14 and the local DB 134b managed by the local DBMS 129b are established and data of the plurality of local DB's are accessed in a consolidated fashion through the center DBMS 14. Fig. 2 is depicted by way of example of a process operation in the center server 101, local A sever 102a and local storage subsystem 104a but the same process may be carried out in the local B server 102b and local storage subsystem 104b.

(1) When a request for consolidated access to data of the plurality of local DB's is made, the replication requesting unit 110 of the center server 101 transmits a replication request for requesting all the local servers 102a and 102b to start replication (step 401).

(2) The DB freeze requesting units 121a and 121b of the local servers 102a and 102b receiving the replication request transmitted in the step 401 make a request to the local DBMS's 129a and 129b for freeze of the local DB's 134a and 134b. "Freeze" referred to herein means a process in which the whole of the update information on buffers in the DBMS's existing on the local servers 102a and 102b is reflected to the local

DB's 134a and 134b and any update subsequently applied to the buffers is inhibited from being reflected upon the local DB's 134 (step 402).

(3) Thereafter, the remote volume split
5 requesting units 122a and 122b of the local servers 102a and 102b request the local storage subsystems 104a and 104b to perform split between local DB 134a and replication local DB 133a and that between local DB 134b and replication local DB 133b. Here, "split"
10 signifies that transfer of the update information to the replication local DB's is stopped (step 403). The updates applied to the local DB's are stored in the local storage subsystems 104.

(4) The local storage subsystems 104a and
15 104b receive the request in the step 403 and the remote volume splitting units 126a and 126b stop transfer of the information for updates applied to the local DB's 134a and 134b to the replication local DB's and perform a process of remote volume split in which the
20 information of updates applied to the local DB's 134a and 134b is stored in the local storage subsystems of their own (step 404).

(5) When the process for remote volume split
in the step 404 is completed, the remote volume split
25 completion notifying units 123a and 123b of the local servers 102a and 102b receive a report to this effect and transfer the fact that the volume split is completed to the center server 101 through the medium

of a remote volume split completion notice (step 405).

(6) Thereafter, the DB freeze release requesting units 124a and 124b of the local servers 102a and 102b make a request to the local DBMS's 129
5 for release of the local DB freeze (step 406).

(7) On the other hand, the center server 101 receives the volume split completion notice in the step 405 and the data consolidating unit 3 executes a process for data consolidation. As will be described
10 later, through the data consolidation process, the contents of the local DB's 134a and 134b are reflected on the shadow images 131a and 131b and after completion of this process, a consolidated access can be ensured by accessing the shadow images 131a and 131b by way of
15 the center DBMS 14 (step 407).

(8) With the process in the step 407 completed, the data consolidation completion notifying unit 112 of the center server 101 informs all of the local servers 102a and 102b that the data consolidation
20 is completed (step 408).

(9) The local servers 102a and 102b receive the notice of data consolidation completion in the step 408 and the remote volume resynchronization requesting units 125a and 125b request the local storage
25 subsystems 104a and 104b to execute remote volume resynchronization for performing resynchronization of the remote volumes. "Resynchronization" referred to herein is a process in which the update information

stored in the local storage subsystems as a result of the aforementioned split process is transferred to the center storage subsystem 103 so that the update information may be reflected upon the replication local
5 DB's 133a and 133b (step 409).

(10) The local storage subsystems 104a and 104b receive the request for remote volume resynchronization and the remote volume resynchronizing units 127a and 127b transfer the update information
10 stored in the local storage subsystems of their own to the center storage subsystem 103 to enable it to perform the resynchronization process for reflecting the update information upon the replication local DB's 133a and 133b. After the resynchronization, the
15 updates applied to the local DB's 134a and 134b are asynchronously transferred to the center storage subsystem 103 and besides reflected on the local DB's 133a and 133b asynchronously by the volume replicating unit 117 (step 410).

20 Fig. 3 is a table for explaining an example of structure of the replication source managing table 2 provided in the center server 101. The shown replication source managing table 2 includes a replication source DB name field 201 indicative of
25 names of replication source servers and a replication completion flag 202 indicating whether a replication between local DB 134a and replication local DB 133a and that between local DB 134b and replication local DB

133b are completed. All flags 202 are initialized with "unfinished" when the replication requesting unit 110 starts the process. In the example shown in Fig. 3, there are records 203a and 203b, with the record 203a
5 indicating that synchronization between the local DB 134a managed by the local DBMS 129a of local A server 102a and the replication local DB 133a is not finished and with the record 203b indicating that synchro-
nization between the local DB 134b managed by the local
10 DBMS 129b of local B server 102b and the replication local DB 133b is not finished.

Fig. 4 is a flowchart for explaining a process executed in the data consolidating unit 3 provided in the center server 101 (the process in the
15 step 407 explained in connection with Fig. 2). The process for data consolidation will now be described with reference to Fig. 4.

(1) When receiving a remote volume split request completion notice from one of the local servers
20 102a and 102b through the LAN 105 representing the first network, the data consolidating unit 3 searches or retrieves, from the replication source managing table 2, a record 203a or 203b corresponding to a name at replication source DB name field 201 on the basis of
25 the name of the local server 102a or 102b which has transmitted the notice and changes the value of replication completion flag field 202 at the corresponding record 203a or 203b to "finished" (steps

301 and 302).

(2) Next, it is checked whether the values of replication completion flag field 202 at all records 203a and 203b stored in the replication source managing
5 table 2 are "finished" in order to check whether notices of split completion, that is, replication completion from all the local servers are received. If even a single "unfinished" record exists, the process is ended to wait for a notice of split completion from
10 another local server (step 303).

(3) In case the values of replication completion flag field 202 are determined to be "finished" through checking in the step 303, a request is made to the center DBMS 14 for freeze of the shadow images 131a and
15 131b managed by the DBMS 14. "Freeze" referred to herein is a process in which all updates on a buffer in the DBMS 14 existing on the center server 101 are reflected to the shadow images 131a and 131b and any update subsequently applied to the buffer is inhibited
20 from being reflected upon the shadow images 131 (step 304).

(4) Thereafter, a volume synchronization request is made to the volume synchronizing unit 115 of center storage subsystem 103. The volume synchronizing
25 unit 115 receiving the request reflects the updates applied to the replication local DB's 133a and 133b from the local storage subsystems 104a and 104b upon the shadow images 131a and 131b. With the reflection

upon the shadow images completed, a request for volume split is made to the volume splitting unit 116 of center storage subsystem 103 to perform split between the shadow image 131a and the replication local DB 133a and that between the shadow image 131b and the replication local DB 133b (steps 305 and 306).

(5) The volume splitting unit 116 receiving the volume split request performs split between each of the shadow images 131 and each of the replication local DB's 133. Thereafter, a request is made to the center DBMS 14 for release of freeze (step 307).

Fig. 5 shows an example of the replication source managing table 2 after synchronization of only the DB at local area A has completed and Fig. 6 shows an example of the replication source managing table 2 after synchronization of the two DB's at local areas A and B.

Next, an example of operation of the data consolidating unit 3 when the two local DB's at local areas A and B shown in Fig. 1 are replicated in the center storage subsystem 103 and replications are accessed from the center server 101 in a consolidated fashion will be described using the examples of replication source managing table 2 shown in Figs. 3, 5 and 6.

Firstly, when the data consolidating unit 3 receives a notice of remote volume split completion from the local A server 102a (step 301), it changes,

through the process of step 302, the replication source managing table, which has already been explained in connection with Fig. 3, in such a manner that the value of replication completion flag field 202 at record 203a
5 indicating whether a replication of the local DB 134a at area A is completed is changed to "finished" as shown in Fig. 5 (step 302). The fact that the record 203a is a record indicative of replication completion of the local DB 134a at area A can be determined from
10 the fact that the value of replication source DB name field 201 corresponding to this record is "DB at area A".

It is subsequently checked through the process of step 303 if the values of replication
15 completion flag field 202 of the table 2 are all "finished". In this phase, the status of table 2 is as shown in Fig. 5, so that the decision result is "No" and the process ends.

Next, when the data consolidating unit 3
20 receives a remote volume split completion notice from the local B server 102b (step 301), it changes through the process of step 302 the replication source managing table 2 shown in Fig. 5 such that the value of replication completion flag field 202 at record 203a
25 indicating whether a replication of the local DB 134b at area B is completed is changed to "finished" as shown in Fig. 6 (step 302).

Next, it is checked through the process of

step 303 if the values of replication completion flag field 202 of table 2 are all "finished". In this phase, the table 2 is conditioned as shown in Fig. 6 and the decision result is "Yes" (step 303).

5 Thereafter, a request is made to the center DBMS 14 for freeze of the center DB's (shadow images 131a and 131b) through the process of step 304. Then the volume synchronizing unit 115 of center storage subsystem 103 is requested to synchronize replication
10 local DB 133a with shadow image 131a and to synchronize replication local DB 133b with shadow image 131b (step 305).

 With the synchronization completed, the volume splitting unit 116 of center storage subsystem
15 103 is requested to split the replication local DB 133a from the shadow image 131a and to split the replication local DB 133b from the shadow image 131b (step 306). Thereafter, a request is made to the center DBMS 14 for release of freeze of the center DB's (step 307).

20 The individual processes according to the embodiment of the invention described so far can be implemented with a process program, which process program can be stored in a recording medium such as HD, DAT, FD, MO, DVD-ROM or CD-ROM and can be presented.

25 As described above, in the embodiment of the invention, starting from a status in which updates are applied asynchronously from the local DB's 134a and 134b to the replication DB's 133a and 133b representing

replications of the local DB's, the local DB's 134a and 134b are first frozen on the basis of a replication request so as to be placed in matched condition and then the local DB 134a is split from the replication
5 local DB 133a and the local DB 134b is split from the replication local DB 133b, so that the replication local DB's 133a and 133b can instantaneously be placed in synchronized condition. On the side of local servers 102a and 102b, the DB freeze is thereafter
10 released (but the volume remains to be split) to permit normal business affairs to proceed and therefore, the freezing period of DB's can be shortened. The center server 101, on the other hand, freezes the center DB's (shadows 131a and 131b) at the time that replications
15 of all of the replication local DB's are completed to carry out the synchronization process.

In this manner, business affairs can be continued before the freeze of center DB's is started. After freeze of the center DB's (shadow images 131a and
20 131b), the process for synchronization between the replication local DB's 133a and 133b and the local shadow images 131a and 131b is carried out but this process replicates only the update bit map and can be ended within a short period of time. (After the
25 synchronization process, the center storage subsystem 103 copies the corresponding records from the replication local DB's 133a and 133b to the local shadow images 131a and 131b.)

According to the embodiment of the invention described previously, update of all local DB's can be reflected at a time and hence, data of the individual local DB's can be accessed in a consolidated fashion without causing the data to be mismatched. Further, a request is made to the local server for resynchronization between the local DB and the replication local DB after completion of synchronization with the shadow image, and the update information stored in the local storage subsystem while the local DB is split from the replication local DB is transmitted to the center storage subsystem, thereby ensuring that the update information can again be reflected upon the replication local DB.

According to the foregoing embodiment of the invention, in the DB system comprising a center server, one or more local servers, a first network for connecting the center server and the local servers, local storage subsystems for storing local DB's managed by the local servers, a center storage subsystem for storing replications of the local DB's and a second network for connecting the center server, center storage subsystem, local servers and local storage subsystems, the center server includes a replication requesting unit, a data consolidating unit, a center DB freeze requesting unit, a center DB freeze release requesting unit, a replication source managing table and a data consolidation completion notifying unit,

each of the local servers includes a local DB freeze
requesting unit, a DB replicating unit, a remote volume
split completion notifying unit and a local DB freeze
release requesting unit, the center storage subsystem
5 stores replications of the local DB's stored in the
local storage subsystems and shadow images of the
replicated local DB's and includes a volume replicating
unit, a volume splitting unit and a volume
synchronizing unit and each of local storage subsystems
10 includes a remote volume replicating unit, a remote
volume splitting unit and a remote volume resynchroniz-
ing unit through a remote DB cooperating scheme,
whereby a plurality of DB's at remote locations can be
accessed instantaneously in a consolidated fashion.

15 As described above, according to the
invention, consolidated access to the plurality of DB's
at remote locations can be ensured instantaneously.

 It should be further understood by those
skilled in the art that although the foregoing
20 description has been made on embodiments of the
invention, the invention is not limited thereto and
various changes and modifications may be made without
departing from the spirit of the invention and the
scope of the appended claims.